Embedding in law of discrete time ARMA processes in continuous time stationary processes

Argimiro Arratia[a,1], Alejandra Cabaña[b,2], Enrique M. Cabaña[c]

[a] *Universitat Politècnica de Catalunya, Spain*
[b] *Universitat Autònoma de Barcelona, Spain*
[c] *Universidad de la República, Montevideo, Uruguay*

Abstract

Given any stationary time series $\{X_n : n \in \mathbf{Z}\}$ satisfying an ARMA(p, q) model for arbitrary p and q with infinitely divisible innovations, we construct a continuous time stationary process $\{x_t : t \in \mathbf{R}\}$ such that the distribution of $\{x_n : n \in \mathbf{Z}\}$, the process sampled at discrete time, coincides with the distribution of $\{X_n\}$. In particular the autocovariance function of $\{x_t\}$ interpolates that of $\{X_n\}$.

Keywords: Discrete-time ARMA, continuous-time ARMA, CARMA, Lévy process, embedding

---

1. Introduction

The description of phenomena that evolve continuously in time is frequently performed by means of series of observations made at equally spaced time intervals. When some regularity is assumed on the conditions under which the observations are made, a discrete-time stationary series might be used as a model for the observations. It is worth noticing that such model can be applied either for equally spaced observations of a stationary phenomenon

Email addresses: argimiro@cs.upc.edu (Argimiro Arratia), acabana@mat.uab.ca (Alejandra Cabaña), ecabana@fing.edu.uy (Enrique M. Cabaña)

[1]Supported by Ministerio de Economía, Industria y Competitividad, Spain, project APCOM [TIN2014-57226-P] and Generalitat de Catalunya, project MACDA [SGR2014-890]

[2]Supported by Ministerio de Economía, Industria y Competitividad, Spain, project MTM2015-69493-R

that evolves in continuous-time (for instance, the concentration of oxygen in blood of a newborn child measured at fixed intervals), or as the result of observing a periodic phenomenon at the same point of successive periods (for instance, the temperature at noon at a given location in the equator). In the first case, a continuous-time model for the stationary process should exist, associated to the discrete-time model which is embedded in it. This poses the problem of finding such a process. The following particular case of this problem has been extensively studied by several authors: Given the stationary process X_t, $t \in \mathbf{Z}$, that satisfies the discrete ARMA(p, q) model, denoted as DARMA(p, q), namely

$$X_t = \sum_{j=1}^{p} \phi_j X_{t-j} + \sum_{k=0}^{q} \theta_k \epsilon_{t-k} \tag{1}$$

where ϵ_t is a white noise with finite variance, the problem of obtaining a stationary process $x_t, t \in \mathbf{R}$, satisfying a continuous version of the DARMA model (denoted CARMA(p, q)) such that when sampled at discrete times has the same autocovariance function as $\{X_t\}$, is known as *embedding* a discrete-time ARMA process in a continuous-time ARMA process.

Brockwell (1995, 2004) summarises the construction of CARMA processes as follows: A CARMA(p, q) process, for $0 \le q < p$, is defined formally as a stationary solution of a continuous analogue of (1), namely, the stochastic differential equation

$$a(\mathcal{D})Y_t = b(\mathcal{D})\mathcal{D}\Lambda_t, \quad t \in \mathbf{R} \tag{2}$$

where $\mathcal{D}$ denotes differentiation with respect to t, Λ is a second-order Lévy process, and

$$\begin{aligned} a(z) &= z^p + a_1 z^{p-1} + \ldots + a_p \\ b(z) &= b_0 + b_1 z + \ldots + b_q z^q \end{aligned}$$

are polynomials of order p and q, respectively. These CARMA processes are linear functions of continuous vector autoregressive (CVAR) Markovian processes.

The embedding problem has been studied by several authors. The works by Chan and Tong (1987), He and Wang (1989), Brockwell (1995) and Brockwell and Brockwell (1999) established embeddings of some DARMA(p, q) processes in continuous ARMA(p, q), for $0 \le q < p$. Huzii (2006) gave necessary and sufficient conditions for a DARMA process to be embedded in a CARMA process. Using the concept of generalised random process of Gel'fand and Vilenkin (1964), Brockwell and Hannig (2010) extended the above definition of CARMA processes to allow for $q \geq p$. However in this case the generalised CARMA process does not exist in the classical sense.

All these approaches to the embedding problem are only concerned with the covariance structure of the processes involved, not with their probability distributions besides the fact that, if the processes are Gaussian, the equality of the first- and second-order moments entails the equality of the probability laws. In general, the discretised version of the CARMA will not necessarily have the same law as the original DARMA.

We propose in this work a different approach to construct for any DARMA (p, q) a continuous stationary *embedding in law*. Our main result is the following:

Theorem 1. *Given the stationary causal DARMA(p, q) process X_t that satisfies (1) with centred infinitely divisible innovations ϵ_t with finite variance, there exists at least one function $L : \mathbf{R}^+ \to \mathbf{R}$ decaying exponentially at infinity and a centred second order Lévy process Λ on $\mathbf{R}$, such that the stationary process $x_t = \int_{-\infty}^{t} L(t-s)d\Lambda(s), t \in \mathbf{R}$, when sampled at times $t \in \mathbf{Z}$, has the same joint law as X_t, $t \in \mathbf{Z}$.*

The function L and the Lévy process Λ are not unique in general. Both must fulfil conditions related to the discrete noise ϵ. The construction that proves this statement is based on the expression of the causal DARMA (1) as an infinite moving average $X_n = \sum_{j=0}^{\infty} b_j \epsilon_{n-j}$ with i.i.d. white noise ϵ_j centred and with finite variance.

A continuous-parameter process $x_t = \int_{-\infty}^{t} L(t-s)d\Lambda(s)$, where Λ and L satisfy the conditions of the statement is stationary, and $(x_n)_{n\in\mathbf{Z}}$ has the same law as $(X_n)_{n\in\mathbf{Z}}$ provided a suitable square integrable function l with domain $[0, 1)$ satisfies $\int_0^1 l(1-s)d\Lambda(s) \sim \epsilon_1$ and $L(s+j) = b_j l(s)$ for $0 \leq s < 1$ and $j = 0, 1, 2 \ldots$.

To describe the family of pairs $\{(\Lambda, l) : \int_0^1 l(1-s)d\Lambda(s) \sim \epsilon\}$ is an open problem, as far as we know, but a necessary and sufficient condition for this family to be nonempty is that ϵ have an infinitely divisible law. In fact, the integral with respect to the Lévy process has an infinitely divisible law, and if ϵ is infinitely divisible, the pair $(\Lambda, 1)$ such that $\Lambda(1) \sim \epsilon_1$ belong to the family (the law of a Lévy process Λ is determined by the law of $\Lambda(1)$). Other pairs that belong to the family are mentioned in Section 3.4.

The details of the construction, that makes use of a similar embedding for vectorial autoregressive (VAR) processes driven by an infinitely divisible white noise, are described below.

The embedding x_t can be applied to analyse the properties of the paths on intervals between observed points, such that the maxima, minima or the properties of the sagittae, that is, the difference between x_t and the (non-stationary) linear interpolation joining the graph of observed values.

The rest of the paper is structured as follows: The following section contains a brief outline of the construction of the embedding. Section 3 describes the technical details of the general construction, and Section 4 discusses the selection of a functional parameter of the embeddings in order to get desirable probabilistic properties, based on the properties of the autocovariances of the embedding. Finally, Appendices A and B supply the detailed computation of the Jordan normal form appearing in Section 3.2.

2. Proposed scheme for the embedding

We consider a family of continuous parameter stationary processes

$$x_t = \int_{-\infty}^{t} L(t-s) d\Lambda_s$$

depending on a centred second-order Lévy process $\Lambda_s, s \in \mathbf{R}$, and on a square integrable function $L(t), t \in \mathbf{R}^+$, and show that given any DARMA $X_t, t \in \mathbf{Z}$, the parameters Λ and L can be selected in order that $\{X_t : t \in \mathbf{Z}\}$ and $\{x_t : t \in \mathbf{Z}\}$ have the same law.

The construction of the embedding is performed in five steps:

(1) Express the scalar DARMA(p, q) as an r dimensional discrete vectorial autoregressive process DVAR(1) where $r = \max\{p, q+1\}$.

(2) Transform the DVAR(1) into a new vectorial process J-DVAR(1) associated to the Jordan canonical form J of the matrix defining the DVAR(1).

(3) Split the J-DVAR(1) into simpler processes associated to each of the Jordan blocks in J, whose dimensions add up to r.

(4) Construct a continuous parameter embedding for each process in the previous step and join them in an embedding J-CVAR(1) for J-DVAR(1). This involves the construction of the Lévy integrator.

(5) Retrace the previous steps from the J-CVAR and obtain a continuous embedding of DARMA.

Next section contains a description of each step.

3. Construction of the embedding

3.1. From DARMA(p,q) to DVAR(1) in state space of dimension $r = \max\{p, q+1\}$

Let $(\epsilon_t)_{t\in\mathbf{Z}}$ denote a standardised white noise (ϵ_k are i.i.d. with $\mathbf{E}\epsilon_1 = 0, \mathbf{E}\epsilon_1^2 = 1$), D the $r \times r$ matrix

$$D = \begin{pmatrix} \phi_1 & \phi_2 & \phi_3 & \dots & \phi_{r-1} & \phi_r \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\eta}_t = \epsilon_t \boldsymbol{u}, \boldsymbol{u} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Then it is well known (Doob, 1944, Sec. 4) (cf. Example 8.3.2. in Brockwell and Davis (2002)) that the stationary causal series satisfying the DARMA(p,q) model in equation (1) can be expressed as the linear function

$$X_t = \boldsymbol{\theta}^{\mathrm{tr}} \boldsymbol{\xi}_t = \sum_{k=0}^{q} \theta_k \xi_{t,k+1}$$

of the DVAR(1)

$$\boldsymbol{\xi}_t = D\boldsymbol{\xi}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\xi}_t = (\xi_{t,1}, \xi_{t,2}, \dots, \xi_{t,r})^{\mathrm{tr}} \tag{3}$$

where $r = \max\{p, q+1\}$, $\phi_j = 0$ for $j > p$, and $\boldsymbol{\theta}^{\mathrm{tr}} = (\theta_0, \theta_1, \dots, \theta_{r-1})$, with $\theta_k = 0$ for $k > q$. (By A^{tr} we denote the transpose of matrix A.)

3.2. From DVAR(1) to J-DVAR(1)

The eigenvalues of the matrix D are the roots ρ_h of the polynomial equation $\rho^r = \sum_{j=1}^{r} \phi_j \rho^{r-j}$, with algebraic multiplicities m_h. In other words, the eigenvalues are the inverses of the roots of the polynomial

$$\phi(z) = 1 - \sum_{j=1}^{p} \phi_j z^j \tag{4}$$

associated to the AR-coefficients of the DARMA(p, q) model, with their algebraic multiplicities, $\phi_p \neq 0$, and, whenever $q \geq p$, also $\rho_0 = 0$ with multiplicity $r - p$.

The space of solutions of each of the equations $D\boldsymbol{v}_h = \rho_h \boldsymbol{v}_h$ has dimension one, and is generated by the eigenvector $\boldsymbol{v}_h = (\rho_h^{-1}, \rho_h^{-2}, \rho_h^{-3}, \ldots, \rho_h^{-r})^{\mathrm{tr}}$ if $\rho_h \neq 0$, and by $\boldsymbol{v}_0 = (0, 0, 0, \ldots, 0, 1)^{\mathrm{tr}}$ for $\rho_0 = 0$.

Let C denote the matrix that carries D to its Jordan canonical form

$$J = C^{-1}DC, \text{ where } J = \begin{pmatrix} J_{\rho_0} & 0 & 0 & \ldots & 0 \\ 0 & J_{\rho_1} & 0 & \ldots & 0 \\ 0 & 0 & J_{\rho_2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & J_{\rho_k} \end{pmatrix}.$$

where $\rho_j \neq \rho_k$ if $j \neq k$. The block $J_{\rho_h} = \rho_h I_{m_h} + I_{1,m_h}$ is associated to the eigenvalue ρ_h with multiplicity m_h, where for each m, I_m is the $m \times m$ identity matrix and $I_{1,m}$ is the $m \times m$ matrix with the first subdiagonal of ones and all other entries equal to zero.

Now the change $\boldsymbol{\xi} = C\boldsymbol{\zeta}$ brings (3) into the canonical form

$$\boldsymbol{\zeta}_t = J\boldsymbol{\zeta}_{t-1} + \epsilon_t C^{-1}\boldsymbol{u}. \tag{5}$$

To obtain the Jordan normal form of a matrix involves delicate numerical computations because the problem to be solved is in general extremely ill conditioned.

However, in our present case the particular form of the matrix D makes possible to write simple expressions for the elements of the matrix C, as described in Theorem 5 included in Appendix A.

Remark 1. The assumption of causality implies that the moduli of $\rho_h, h = 1, \ldots, k$, are smaller than one, and hence $\lim_{n\to\infty} D^n = 0$ or $\lim_{n\to\infty} J^n = 0$. By recursively applying equation (3) one gets for $m = 1, 2, \ldots$

$$\boldsymbol{\xi}_t = D^m \boldsymbol{\xi}_{t-m} + \sum_{n=0}^{m-1} D^n \boldsymbol{\eta}_{t-n} \tag{6}$$

or

$$\boldsymbol{\xi}_t = \sum_{n=0}^{\infty} D^n \boldsymbol{\eta}_{t-n} \tag{7}$$

Since all eigenvalues of D have moduli smaller than one, D^n tends to zero exponentially as n goes to infinity and hence the series in the right-hand member of (7) converges in the space of r-dimensional random vectors with the quadratic norm $\sqrt{\mathbf{E}\|\boldsymbol{\eta}\|^2}$ ($\|\boldsymbol{\eta}\|$ denotes the Euclidian norm). Analogous equations for $\boldsymbol{\zeta}_t$ are obtained from (5).

From (6) follows by multiplying by $\boldsymbol{\xi}_{t-m}^{\mathrm{tr}}$ and taking expectations that the covariance $\mathbf{Cov}(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-m}) = D^m \mathbf{Var}\boldsymbol{\xi}_{t-m} = D^m \mathbf{Var}\boldsymbol{\xi}_t$ only depends on the variance of the noise, but not on its probability law, and this is applied in He and Wang (1989) to show that the DVAR has a (Gaussian) continuous embedding with the same covariances (*but only exceptionally the same distribution!*). Their embedding is a Markovian solution of a stochastic differential vectorial equation.

Equation (7) shows that the distribution of the VAR does depend on the distribution of the noise, and also the nature of that dependence. The stationary embedding we are proposing is not Markovian, but sampled on integer times, not only has the same covariances as the original DVAR but also the same probability distributions, and if the original discrete noise were known, would even coincide a.s. with the DVAR for integer times.

3.3. Splitting J-DVAR(1) into elementary equations and processing each one individually

The vectorial equation (5) with $\boldsymbol{\zeta}_t = \begin{pmatrix} \boldsymbol{\zeta}_{t,\rho_0} \\ \boldsymbol{\zeta}_{t,\rho_1} \\ \vdots \\ \boldsymbol{\zeta}_{t,\rho_k} \end{pmatrix}$ and $C^{-1}\boldsymbol{u} = \begin{pmatrix} c_{\rho_0} \\ c_{\rho_1} \\ \vdots \\ c_{\rho_k} \end{pmatrix}$ partitioned in segments of lengths $m_0, m_1, \ldots, m_k$ is equivalent to the $k+1$ canonical equations to be treated separately

$$\boldsymbol{\zeta}_{t,\rho_h,} = J_{\rho_h} \boldsymbol{\zeta}_{t-1,\rho_h} + \epsilon_t c_{\rho_h}. \tag{8}$$

For imaginary eigenvalues, the equations obtained and their solutions are also imaginary but, assuming that the original DARMA is real, they will be finally combined to obtain a real process (see equation (17) below).

Let us write generically each of the equations in (8) as

$$\boldsymbol{\zeta}_{t,\rho} = J_\rho \boldsymbol{\zeta}_{t-1,\rho} + \epsilon_t c_\rho, \quad J_\rho = \rho I_m + I_{1,m}. \tag{9}$$

The iterated application of (9) leads to write $\zeta_{t,\rho}$ as the sum of the series

$$\zeta_{t,\rho} = \sum_{n=0}^{\infty} \epsilon_{t-n} J_\rho^n c_\rho \tag{10}$$

as noticed in Remark 1. The sum converges because, since $I_{1,m}^m = 0$,

$$J_\rho^n = (\rho I_m + I_{1,m})^n = \sum_{j=0}^{n \wedge (m-1)} \binom{n}{j} \rho^{n-j} I_{1,m}^j$$

has its components uniformly bounded by $mn^m \rho^{n-m+1}$ that, because $|\rho| < 1$, converges exponentially to zero, as n goes to infinity.

In order to extend the domain of $\zeta_{t,\rho}$ to all $t \in \mathbf{R}$, we introduce a representation of the innovations ϵ_t by means of integrals of some function $l : [0,1) \to \mathbf{R}$ with respect to a Lévy process

$$\epsilon_t = \int_{t-1}^{t} l(t-s) d\Lambda(s). \tag{11}$$

The existence of such representation is discussed in Section 3.4.

Then replace the series by an integral of a vector function $\boldsymbol{L}_\rho$ with respect to the Lévy process

$$\zeta_{t,\rho} = \int_{-\infty}^{t} \boldsymbol{L}_\rho(t-s) d\Lambda(s) \tag{12}$$

and impose that it satisfies equation (9).

By using the notation $\mathcal{S}\boldsymbol{L}_\rho(t) = \boldsymbol{L}_\rho(t+1)$ to denote the unitary shift, this amounts to impose that

$$\int_{-\infty}^{t} \boldsymbol{L}_\rho(t-s) d\Lambda(s) = \int_{-\infty}^{t-1} \mathcal{S}\boldsymbol{L}_\rho(t-1-s) d\Lambda(s) + \int_{t-1}^{t} \boldsymbol{L}_\rho(t-s) d\Lambda(s)$$

be equal to

$$J_\rho \int_{-\infty}^{t-1} \boldsymbol{L}_\rho(t-1-s) d\Lambda(s) + \epsilon_t c_\rho.$$

The equality holds termwise provided

(i) $\boldsymbol{L}_\rho(s) = c_\rho l(s)$ for $0 \le s < 1$ and

(ii) $\mathcal{S}\boldsymbol{L} = J_\rho \boldsymbol{L}$.

Starting from (i), the repeated application of (ii) permits to compute recursively $\boldsymbol{L}_\rho$ on each interval between consecutive integers: For each integer

$n > 0$ and $t \in [0,1)$, $\boldsymbol{L}_\rho(n+t) = \boldsymbol{\mathcal{S}L}_\rho(n-1+t) = J_\rho \boldsymbol{L}_\rho(n-1+t) = \ldots = J_\rho^n \boldsymbol{L}_\rho(t) = J_\rho^n c_\rho l(t)$, and therefore

$$\boldsymbol{L}_\rho(t) = J_\rho^{[t]} c_\rho l(\mathrm{frac}(t)) \tag{13}$$

with $[t]$ denoting the integer part of t and $\mathrm{frac}(t) = t - [t]$ its fractional part.

For further convenience, we extend the domain of the function l to all $\mathbf{R}$ and define $l(s) = 0$ for $s \notin [0,1)$.

3.4. Choosing the function l and the Lévy integrator Λ

Assume without loss of generality that $\mathbf{Var}\epsilon_t = \int_0^1 |l(s)|^2 ds = 1$ and $\mathbf{Var}\Lambda(1) = 1$. Equation (11) establishes a relationship between ϵ_t, l and Λ.

A necessary condition for (11) to be satisfied regardless of the choice of l is that the distribution of the noise must be infinitely divisible. But this condition is not sufficient in general: Despite Gaussian noises can be represented with any integrand l with norm one and Λ a Wiener process, integrands $l \neq 1$ limit the family of noises that admit the representation $\epsilon_t = \displaystyle\int_0^1 l(1-s) d\Lambda(s+t-1)$. For instance, when $r = 1$, and $l(s)$ is proportional to $\exp(-\kappa s)$, one obtains the Ornstein-Uhlenbeck classic CARMA interpolation of an AR(1), or DARMA(1,0) (we summarise this construction in Example 2 at the end); if ϵ_t only assumes integer values (such as $\epsilon_t \sim \mathrm{Poisson}(1) - 1$) that particular non constant continuous l cannot ensure that the sum of the jumps of Λ times the values of l at the jumps be integer.

Equation (11) implies that the characteristic functions χ_ϵ and $\chi_{\Lambda(1)}$ of ϵ and $\Lambda(1)$ must satisfy

$$\log \chi_\epsilon(z) = \int_0^1 \log \chi_{\Lambda(1)}(zl(1-s)) ds \tag{14}$$

which in particular holds trivially for $l = 1$ and $\Lambda(1) = \epsilon$, and also for $l(s) = \frac{1}{\sqrt{a}} \mathbf{1}_{\{0 < s \leq a\}}$ and $\Lambda(a) = \epsilon\sqrt{a}$ (notice that the infinitely divisible law of $\Lambda(a)$ for any $a \neq 0$ determines the laws of the Lévy process Λ. These examples lead to different solutions for the embedding, that can be applied to any infinitely divisible noise. Other selections of l depending of the probability distribution of the noise might also be used. We return to this lack of uniqueness in Section 4.1, where an optimality criterion leads to select $l = 1$ in order to maximise the covariances of the embedding.

3.5. Joining the solutions corresponding to each Jordan block and retracing steps

The continuous parameter embedding of (5), that we denote by the same symbol $\boldsymbol{\zeta}_t$ is composed by the juxtaposition of the processes

$$\boldsymbol{\zeta}_{t,\rho_h} = \int_{-\infty}^{t} \boldsymbol{L}_{\rho_h}(t-s)d\Lambda(s)$$

so that, for each selection of the function l, $\boldsymbol{L}_{\rho_h}(t) = J_{\rho_h}^{[t]} c_{\rho_h} l(\mathrm{frac}(t))$, and the matching Lévy process Λ, the continuous parameter stationary process $x_t = \boldsymbol{\theta}^{\mathrm{tr}} C\boldsymbol{\zeta}_t$ is an embedding for X_t.

Let us denote by $C_{\rho_h}, h = 0, 1, \ldots, k$, each of the matrix blocks of size $r \times m_h$ that compose the matrix $C = (C_{\rho_0}, C_{\rho_1}, C_{\rho_2}, \ldots, C_{\rho_k})$. Therefore, with these notations, the embedding x_t can be written as

$$x_t = \theta^{\mathrm{tr}} \sum_{h=0}^{k} C_{\rho_h} \int_{-\infty}^{t} J_{\rho_h}^{[t-s]} c_{\rho_h} l(\mathrm{frac}(t-s))d\Lambda(s). \tag{15}$$

This proves Theorem 1 with

$$L(t) = \theta^{\mathrm{tr}} \sum_{h=0}^{k} C_{\rho_h} J_{\rho_h}^{[t]} c_{\rho_h} l(\mathrm{frac}(t)). \tag{16}$$

When ρ_h is imaginary, also its conjugate $\bar{\rho}_h$ is an eigenvalue with the same multiplicity, and the sum $C_{\rho_h} J_{\rho_h}^{[t-s]} c_{\rho_h} + C_{\bar{\rho}_h} J_{\bar{\rho}_h}^{[t-s]} c_{\bar{\rho}_h}$ is real, thus contributing jointly to (15) with a real term.

For further convenience, let us replace t by $n+t$ with integer n and $0 \le t < 1$ and write the integral on $(-\infty, n+t]$ as the sum of the integrals on the intervals $I_m^-(t) = (m-1+t, m]$ and $I_m^+(t) = (m, m+t]$ for integer $m \le n$. For $s \in I_m^-(t), [n+t-s] = n-m, [-s] = -m$, $\mathrm{frac}(n+t-s) = t+m-s$ and $\mathrm{frac}(-s) = -s+m$. For $s \in I_m^+(t), [n+t-s] = n-m, [-s] = -m-1$, $\mathrm{frac}(n+t-s) = t+m-s$ and $\mathrm{frac}(-s) = -s+m+1$.

Now introduce the abbreviation

$$a_{m,\rho_h} = \theta^{\mathrm{tr}} C_{\rho_h} J_{\rho_h}^{m} c_{\rho_h},$$

and the continuous embedding $\epsilon_t = \int_{t-1}^{t} l(t-s)d\Lambda(s)$ of the innovations process, thus

$$x_{n+t} = \sum_{h=0}^{k} \sum_{m\le n} a_{n-m,\rho_h} \int_{I_m^-(t)\cup I_m^+(t)} l(t+m-s)d\Lambda(s) = \sum_{h=0}^{k} \sum_{m\le n} a_{n-m,\rho_h} \epsilon_{m+t}. \tag{17}$$

In particular, for $t = 0$,

$$x_n = \sum_{h=0}^{k} \sum_{m \leq n} a_{n-m,\rho_h} \epsilon_m = \sum_{m \leq n} A_{n-m} \epsilon_m \text{ with } A_m = \sum_{h=1}^{k} a_{m,\rho_h} \tag{18}$$

reproduces the original X_n. The sum of the terms a_{m,ρ_h} and $a_{m,\bar{\rho}_h} = \bar{a}_{m,\rho_h}$ is real and hence so is A_m.

4. On the covariances of the embedding.

In order to write the following statement, we introduce the notation $\mathcal{S}^t$ for the fractional shift $\mathcal{S}^t f(s) = f(s+t)$ and use the inner product notation $\langle f, g\rangle = \int f(s)\bar{g}(s)ds$.

Theorem 2. *For nonnegative integer n and $0 \leq t < 1$, the covariance $\gamma_{n+t} = \mathbf{E}x_{n+t}x_0$ is related to γ_n and γ_{n+1} by the linear combination*

$$\gamma_{n+t} = \gamma_n \langle \mathcal{S}^{-t} l, l\rangle + \gamma_{n+1} \langle \mathcal{S}^{1-t} l, l\rangle. \tag{19}$$

.

Proof. From (17,18) we have

$$x_{n+t} = \sum_{m \leq n} A_{n-m} \left(\int_{I_m^-(t) \cup I_m^+(t)} l(t+m-s) d\Lambda(s) \right) \tag{20}$$

and

$$\begin{aligned} x_0 &= \sum_{m \leq 0} A_{-m} \left(\int_{m-1}^{m} l(m-s) d\Lambda(s) \right) \\ &= \sum_{m \leq 0} A_{-m} \left(\int_{I_{m-1}^+(t) \cup I_m^-(t)} l(m-s) d\Lambda(s) \right). \end{aligned}$$

Now assume $n \geq 0$ and apply the independence of increments of Λ to get the covariance

$$\gamma_{n+t} = \mathbf{E}x_{n+t}x_0 = \sum_{m \leq 0} A_{n-m} A_{-m} \int_{I_m^-(t)} l(t+m-s) l(m-s) ds$$

$$+\sum_{m<0} A_{n-m}A_{-m-1}\int_{I^+_m(t)} l(t+m-s)l(m+1-s)ds.$$

The change of variables $s = t+m-y$ leads to

$$\gamma_{n+t} = \sum_{m\leq 0} A_{n-m}A_{-m}\int_t^1 l(y)l(y-t)dy$$

$$+\sum_{m<0} A_{n-m}A_{-m-1}\int_0^t l(y)l(y-t+1)dy.$$

$$= \sum_{m\leq 0} A_{n-m}A_{-m}\langle \mathcal{S}^{-t}l, l\rangle + \sum_{m<0} A_{n-m}A_{-m-1}\langle \mathcal{S}^{1-t}l, l\rangle.$$

In particular, for $t = 0$ (respectively $t = 1$), the first inner product is one (respectively zero) and the second is zero (respectively one), so that the covariances of the original DARMA are obtained:

$$\mathbf{E}X_nX_0 = \sum_{m\leq 0} A_{n-m}A_{-m} = \gamma_n, \quad \mathbf{E}X_{n+1}X_0 = \sum_{m<0} A_{n-m}A_{-m-1} = \gamma_{n+1}.$$

and (19) follows.

□

Remark 2. *The simplest result occurs when $l = 1$, because in that case $\langle \mathcal{S}^{-t}l, l\rangle = 1-t$, $\langle \mathcal{S}^{1-t}l, l\rangle = t$ and γ_{n+t} is just a convex combination of γ_n and γ_{n+1}. We show next that this l is solution of an optimisation criterion.*

4.1. Choosing l for maximising the integrated covariance

A completely different behaviour of the covariances occurs when the support of l is a short sub-interval of $[0, 1]$. In that case for t and $1-t$ larger than the length of the support, both coefficients in (19) vanish and the resulting embedding is uncorrelated with (and also independent of) the original process.

Such behaviour can be prevented by requiring that the integrated covariance $\int_0^1 \gamma_t dt$ (or $\int_{-\infty}^{\infty} \gamma_t dt$) be maximum, which is equivalent to require that the integrated variance of the increments $\int_0^1 \mathbf{Var}(x_{n+t}-x_n)dt = 2\int_0^1(\gamma_0-\gamma_t)dt$ be minimum. These comments motivate next theorem.

Theorem 3. *The integrated covariance $\int_0^1 \mathbf{E}x_tx_0dt$ of the embedding x_t constructed as indicated in Section 3 is maximum for $l = 1$ or $l = -1$.*

Proof. Integrate (19) on $[0,1]$ to get

$$\int_0^1 \gamma(t)dt = \int_0^1 (\gamma_0\langle \mathcal{S}^{-t} l, l\rangle + \gamma_1 \langle \mathcal{S}^{1-t} l, l\rangle)dt$$

$$= \gamma_0 \int_0^1 \left(\int_t^1 l(s-t)l(s)ds\right) dt + \gamma_1 \int_0^1 \left(\int_0^t l(s+1-t)l(s)ds\right) dt$$

$$= \gamma_0 \int_0^1 \left(\int_0^s l(s-t)dt\right) l(s)ds + \gamma_1 \int_0^1 \left(\int_s^1 l(s+1-t)dt\right) l(s)ds.$$

Now introduce the function $\mathcal{L}(s) = \int_0^s l(t)dt$ and replace $\int_0^s l(s-t)dt = \mathcal{L}(s)$ and $\int_s^1 l(s+1-t)dt = \mathcal{L}(1) - \mathcal{L}(s)$ in the previous expression, so that

$$\int_0^1 \gamma(t)dt = \gamma_0 \int_0^1 \mathcal{L}(s)l(s)ds + \gamma_1 \int_0^1 (\mathcal{L}(1) - \mathcal{L}(s))l(s)ds$$

$$= \gamma_0 \frac{\mathcal{L}^2(1)}{2} + \gamma_1 \mathcal{L}^2(1) - \gamma_1 \frac{\mathcal{L}^2(1)}{2} = \frac{1}{2}(\gamma_0 + \gamma_1)\mathcal{L}^2(1).$$

Since $\gamma_0 + \gamma_1$ is nonnegative, this expression is maximised when $|\mathcal{L}(1)| = |\langle l, 1\rangle|$ attains its maximum. This occurs when l is proportional to 1, that is, for $l = 1$ or $l = -1$ because of the normalisation $\langle l, l\rangle = 1$. □

Corollary 3.1. *The integrated variance of the increment* $\int_0^1 \mathbf{Var}(x_t - x_0)dt$ *is minimum for* $l = 1$ *or* $l = -1$.

Proof. Since $\mathbf{Var}(x_t - x_0) = 2(\gamma_0 - \gamma_t)$, the result follows immediately from the Theorem. □

Remark 3. *The substitution of* $-l$ *for* l *produces a change of sign in both* L_ρ *and* Λ, *so that the resulting embedding* x *is the same.*

4.2. Covariances of the sagittae

The differences between the embedding and the polygonal interpolation of the original DARMA (i.e. the *sagittae*) are

$$S_{n,t} := x_{n+t} - t x_{n+1} - (1-t)x_n, \;\; t \in [0,1] \tag{21}$$

and the covariance between this process and the values of the DARMA are

$$\Gamma_{k,t} := \mathbf{E} S_{n,t} x_{n-k} = \gamma_{k+t} - t\gamma_{k+1} - (1-t)\gamma_k$$

$$= (\langle \mathcal{S}^{-t}l, l\rangle - (1-t))\gamma_k + (\langle \mathcal{S}^{1-t}l, l\rangle - t)\gamma_{k+1} \tag{22}$$

independent of n because x_t is stationary. The covariance of two such differences is

$$\begin{aligned}
\mathbf{E}S_{n,t}S_{n-k,s} &= \mathbf{E}S_{n,t}(x_{n-k+s} - sx_{n-k+1} - (1-s)x_{n-k}) \\
&= \mathbf{E}S_{n,t}x_{n-k+s} - s\Gamma_{k-1,t} - (1-s)\Gamma_{k,t} \\
&= \mathbf{E}(x_{n+t} - tx_{n+1} - (1-t)x_n)x_{n-k+s} - s\Gamma_{k-1,t} - (1-s)\Gamma_{k,t} \\
&= \gamma_{k+t-s} - t\gamma_{k+1-s} - (1-t)\gamma_{k-s} - s\Gamma_{k-1,t} - (1-s)\Gamma_{k,t}.
\end{aligned}$$

Let us assume $s \le t$, so that $\gamma_{k+t-s} = \gamma_k\langle \mathcal{S}^{-(t-s)}, l\rangle + \gamma_{k+1}\langle \mathcal{S}^{1-(t-s)}, l\rangle$, hence

$$\mathbf{E}S_{n,t}S_{n-k,s} = \gamma_k\langle \mathcal{S}^{-(t-s)}l, l\rangle + \gamma_{k+1}\langle \mathcal{S}^{1-(t-s)}l, l\rangle - t(\gamma_k\langle \mathcal{S}^{1-s}l, l\rangle + \gamma_{k+1}\langle \mathcal{S}^{-s}l, l\rangle)$$

$$-(1-t)(\gamma_k\langle \mathcal{S}^{-s}l, l\rangle + \gamma_{k-1}\langle \mathcal{S}^{1-s}l, l\rangle) - s\Gamma_{k-1,t} - (1-s)\Gamma_{k,t}.$$

Now replace in the preceding formula the following expressions obtained from (22):

$$\gamma_k\langle \mathcal{S}^{-(t-s)}, l\rangle + \gamma_{k+1}\langle \mathcal{S}^{1-(t-s)}, l\rangle = \Gamma_{k,t-s} + (t-s)\gamma_{k+1} + (1-t+s)\gamma_k$$

$$\gamma_k\langle \mathcal{S}^{1-s}l, l\rangle + \gamma_{k+1}\langle \mathcal{S}^{-s}l, l\rangle = \Gamma_{k,1-s} + (1-s)\gamma_{k+1} + s\gamma_k$$

$$\gamma_k\langle \mathcal{S}^{-s}l, l\rangle + \gamma_{k-1}\langle \mathcal{S}^{1-s}l, l\rangle = \Gamma_{k-1,1-s} + (1-s)\gamma_k + s\gamma_{k-1}$$

to get

$$\mathbf{E}S_{n,t}S_{n-k,s} = \Gamma_{k,t-s} + (t-s)\gamma_{k+1} + (1-t+s)\gamma_k - t(\Gamma_{k,1-s} + (1-s)\gamma_{k+1} + s\gamma_k)$$

$$-(1-t)(\Gamma_{k-1,1-s} + (1-s)\gamma_k + s\gamma_{k-1}) - s\Gamma_{k-1,t} - (1-s)\Gamma_{k,t}$$

$$= (t-s)\gamma_{k+1} + (1-t+s)\gamma_k - t((1-s)\gamma_{k+1} + s\gamma_k)$$

$$-(1-t)((1-s)\gamma_k + s\gamma_{k-1}) + \Gamma_{k,t-s} - t\Gamma_{k,1-s} - s\Gamma_{k-1,t} - (1-t)\Gamma_{k-1,1-s} - (1-s)\Gamma_{k,t}$$

$$= s(1-t)(2\gamma_k - \gamma_{k+1} - \gamma_{k-1}) + \Gamma_{k,t-s} - t\Gamma_{k,1-s} - s\Gamma_{k-1,t} - (1-t)\Gamma_{k-1,1-s} - (1-s)\Gamma_{k,t}$$

The general expression is obtained by replacing s and t by $s \wedge t$ and $s \vee t$ respectively.

Again, by choosing $l = 1$, a substantially reduced expression is obtained as summarised in the following statement, because, for all k and t, $\Gamma_{k,t} = 0$.

Theorem 4. *For intervals between consecutive integers $n, n+1$, the sagittae $S_{n,t} = x_{n+t} - (1-t)x_n - tx_{n+1}, 0 \leq t \leq 1$ of the embedding x_t constructed as indicated in Section 3 with the same integrand $l = 1$ have the following properties:*

(i) they are uncorrelated with the values of the original DARMA,

(ii) the autocovariances between two sagittae

$$\mathbf{E} S_{n,t} S_{n-k,s} = s \wedge t(1 - s \vee t)(2\gamma_k - \gamma_{k+1} - \gamma_{k-1})$$

are proportional to the covariances of a Brownian bridge b (that is, the sagitta $b_{n+t} - (1-t)w_n - tw_{n+1}$ of a Wiener process w corresponding to an interval of unit length) and

(iii) the proportionality coefficient is equal to $\mathbf{Cov}(x_n - x_{n-1}, x_{n-k} - x_{n-k-1})$.

Corollary 4.1. *When the original process is Gaussian, $(\epsilon_t)_{t \in \mathbf{Z}}$ are i.i.d. standard normal, Λ is a standard Wiener process and $l = 1$, the embedding can be constructed by adding to each side of the polygonal joining the vertices (n, X_n) and $(n+1, X_{n+1})$, $n \in \mathbf{z}$, Brownian bridges taken from a sequence $(b_{n,t})_{n \in \mathbf{Z}}$ independent of $(X_n)_{n \in \mathbf{Z}}$ that satisfies the same DARMA model as $X_n - X_{n-1}$, with standard independent Brownian bridges $(\beta_n)_{n \in \mathbf{Z}}$ as noise, namely,*

$$b_{n,\cdot} = \sum_{j=1}^{p} \phi_j b_{n-j,\cdot} + \sum_{i=0}^{q} \theta_i (\beta_{n-i} - \beta_{n-1-i}).$$

Example 1. *Assume that the stationary series X_t, $t \in \mathbf{Z}$, satisfies the AR(1) model $X_n = \rho X_{n-1} + \epsilon_n$ where ϵ_n is a Gaussian white noise; that the stationary sequence of processes B_n with domain $[0,1]$, satisfies the model $B_n = \rho B_{n-1} + \beta_n$ where β_n is a sequence of Brownian bridges independent of the noise ϵ. Then the process*

$$\zeta_t = X_{[t]} + (t - [t])(X_{[t]+1} - X_{[t]}) + B_{[t]}(t - [t]), \quad t \in \mathbf{R}$$

is a stationary interpolation of X_n, $n \in \mathbf{Z}$.

Example 2. *To obtain a $CARMA(1,0)$ process as a solution of (2), where to embed a AR(1) process, we need to take l different from 1. In fact, it suffices to take $l(s) = \rho^s$. The resulting*

$$\zeta_t = \int_{-\infty}^{t} \rho^{t-s} d\Lambda(s) = \int_{-\infty}^{t} \mathrm{e}^{-\kappa(t-s)} d\Lambda(s)$$

with $\kappa = -\log\rho$ is an Ornstein-Uhlenbeck process, with the Markov property. Negative values of ρ lead to complex processes, and there is no solution for $\rho = 0$. The admissible probability law of the noise ϵ_t is limited to the ones that can be represented as integrals $\int_0^1 \rho^{1-s} d\Lambda_s$.

Appendices

A. Computation of the matrix C

Let us remind that the notations C_{ρ_h} and J_{ρ_h} for the $r \times m_h$ column blocks of C and $m_h \times m_h$ diagonal blocks of J have already been introduced. These matrices are characterised by the relations

$$DC_{\rho_h} = C_{\rho_h} J_{\rho_h}.$$

Therefore, if $c_{\cdot,j}^{(h)}$ is the j-th column of C_{ρ_h}, then, because $J_{\rho_h} = \rho_h I + I_{m,1}$, the equations

$$(D - \rho_h I) c_{\cdot,j}^{(h)} = c_{\cdot,j+1}^{(h)} \tag{23}$$

must hold for $j = 1, 2, \ldots, m_h$ with $c_{\cdot,m+1}^{(h)} = 0$.

Let $M_{i,j}$ be the matrix of zeros, except for the element in row i and column j that is equal to one. Then the procedure applied to reduce $D - \rho I$ to an essentially super-diagonal matrix can be summarised as follows:

$$D - \rho I = (I + (\phi_1 - \rho) M_{1,2}) \begin{pmatrix} 0 & \phi_2 + \rho\phi_1 - \rho^2 & \phi_3 & \ldots & \phi_{r-1} & \phi_r \\ 1 & -\rho & 0 & \ldots & 0 & 0 \\ 0 & 1 & -\rho & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 1 & -\rho \end{pmatrix}$$

$$= (I + (\phi_1 - \rho) M_{1,2})(I + (\phi_2 + \rho\phi_1 - \rho^2) M_{1,3})$$

$$\times \begin{pmatrix} 0 & 0 & \phi_3 + \rho\phi_2 + \rho^2\phi_1 - \rho^3 & \ldots & \phi_{r-1} & \phi_r \\ 1 & -\rho & 0 & \ldots & 0 & 0 \\ 0 & 1 & -\rho & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & 1 & -\rho \end{pmatrix} = \ldots$$

$$= (I + (\phi_1 - \rho) M_{1,2})(I + (\phi_2 + \rho\phi_1 - \rho^2) M_{1,3})(I + (\phi_3 + \rho\phi_2 + \rho^2\phi_1 - \rho^3) M_{1,4}) \ldots$$
$$\times (I + (\phi_{r-1} + \rho\phi_{r-2} + \ldots + \rho^{r-2}\phi_1 - \rho^{r-1}) M_{1,r})$$

$$\times \begin{pmatrix} 0 & 0 & \dots & 0 & \phi_r + \rho\phi_{r-1} + \dots + \rho^{r-1}\phi_1 - \rho^r \\ 1 & -\rho & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -\rho \end{pmatrix}.$$

Notice that the last component of the first row of the matrix just obtained is the polynomial $\psi(\rho)$ defined by

$$\psi(\rho) := -\rho^r \phi(1/\rho) \tag{24}$$

for $\rho \neq 0$, where ϕ is the polynomial in (4). Therefore its roots are the eigenvalues of D with their respective multiplicities, including the null eigenvalue if r is greater than p.

Also notice that all products $M_{i,j}M_{i',j'}$ with $i < j$ and $i' < j'$ vanish, and hence for any a and $i < j$, $(I + aM_{i,j})^{-1} = I - aM_{i,j}$.

These observations lead to express the conditions (23) as

$$\begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & -\rho_h & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -\rho_h \end{pmatrix} c^{(h)}_{\cdot,j}$$

$$= \begin{pmatrix} 1 & \rho_h - \phi_1 & \rho_h^2 - \rho_h\phi_1 - \phi_2 & \rho_h^3 - \rho_h^2\phi_1 - \rho_h\phi_2 - \phi_3 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots \end{pmatrix} c^{(h)}_{\cdot,j+1}$$

The scalar equation corresponding to the equality of the first components requires to establish that

$$\begin{aligned} & c^{(h)}_{1,j+1} + (\rho_h - \phi_1)c^{(h)}_{2,j+1} + (\rho_h^2 - \rho_h\phi_1 - \phi_2)c^{(h)}_{3,j+1} + \dots \\ & \quad +(\rho_h^{r-1} - \rho_h^{r-2}\phi_1 - \dots - \rho_h\phi_{r-1} - \phi_{r-1})c^{(h)}_{r,j+1} \\ & \quad = \sum_{i=1}^{r} c^{(h)}_{i,j+1}\rho_h^{i-1} - \sum_{j'=1}^{r-1} \phi_{j'} \sum_{i=j'+1}^{r} c^{(h)}_{i,j+1}\rho_h^{i-j'-1} \end{aligned} \tag{25}$$

vanishes for $j = 1, 2, \dots, m_h$. This been granted, the columns of C_{ρ_h} are obtained recursively from right to left and for each column from the last component to be chosen arbitrarily up to the first one, by means of the equations

$$c_{i,j}^{(h)} - \rho_h c_{i+1,j}^{(h)} = c_{i+1,j+1}^{(h)}, \tag{26}$$

for $j = m_h, m_h - 1, \ldots, 1$ and for $i = r-1, r-2, \ldots, 1$.

These equations can be used to obtain a simple algorithm to compute numerically the matrix C and, together with (25), also to prove the closed expression stated in next Theorem.

Theorem 5. *If ρ_h is a root of ψ with multiplicity m_h, the element at the i-th row and j-th column of the $r \times m_h$ matrix C_{ρ_h} corresponding to that root can be chosen as*

$$c_{i,j}^{(h)} = \binom{r-i}{m_h - j} \rho_h^{r-m_h+j-i} \tag{27}$$

with the understanding that $\binom{a}{b}$ for $a < b$ means zero.

Let $\mathcal{D}$ denote the derivative with respect to ρ_h. Then an alternative expression for (27) is

$$c_{i,j}^{(h)} = \frac{\mathcal{D}^{m_h - j} \rho_h^{r-i}}{(m_h - j)!}. \tag{28}$$

Remark 4. *A matrix satisfying (25) and (26) is not uniquely determined. The one with components (27) or (28) is just one particular solution.*

The following lemma prepares the proof of Theorem 5.

Lemma 1. *The equalities*

$$\sum_{i=j+1}^{r-m} \binom{r-i}{m} = \binom{r-j}{m+1}$$

hold for $m = 0, 1, \ldots, r-1$ and $j = 0, 1, \ldots, r-m-1$.

Proof. Apply the so called Stifel Formula that describes the law of construction of Tartaglia's triangle to write $\binom{r-i}{m} = \binom{r-i+1}{m+1} - \binom{r-i}{m+1}$ that holds for $i < r-m$, and also for $i = r-m$ by convening $\binom{m}{m+1} = 0$. After the substitution of the right-hand member of these equalities for each term of the sum in the statement, one obtains the telescopic expression

$$\sum_{i=j+1}^{r-m} \binom{r-i}{m} = \sum_{i=j+1}^{r-m} \left[\binom{r-i+1}{m+1} - \binom{r-i}{m+1} \right]$$

that immediately reduces to the required result. □

Proof of Theorem 5: To establish that equation (26) holds requires to show that

$$\binom{r-i}{m_h-j}\rho_h^{r-m_h+j-i}-\rho_h\binom{r-i-1}{m_h-j}\rho_h^{r-m_h+j-i-1}=\binom{r-i-1}{m_h-j-1}\rho_h^{r-m_h+j-i}$$

and this equality is immediate from the Stifel Formula.

As for verifying that (25) vanishes, notice that for $j=m_h$ it holds trivially because $c^{(h)}_{\cdot,m+1}$ is zero. Then, with m instead of m_h-j-1, replace (27) in (25) to get

$$\begin{aligned}&\sum_{i=1}^{r}\binom{r-i}{m}\rho_h^{r-m-1}-\sum_{j'=1}^{r-1}\phi_{j'}\sum_{i=j'+1}^{r}\binom{r-i}{m}\rho_h^{r-m-j'-1}\\ &=\sum_{i=1}^{r-m}\binom{r-i}{m}\rho_h^{r-m-1}-\sum_{j'=1}^{r-m-1}\phi_{j'}\sum_{i=j'+1}^{r-m}\binom{r-i}{m}\rho_h^{r-m-j'-1}\end{aligned}\qquad(29)$$

and it remains to show that this expression vanishes for each h and $m=0,1,\ldots,m_h-2$.

From Lemma 1, $\sum_{i=1}^{r-m}\binom{r-i}{m}=\binom{r}{m+1}$ and $\sum_{i=j'+1}^{r-m}\binom{r-i}{m}=\binom{r-j'}{m+1}$ so that (29) is equal to

$$\binom{r}{m+1}\rho_h^{r-m-1}-\sum_{j'=1}^{r-m-1}\phi_{j'}\binom{r-j'}{m+1}\rho_h^{r-m-j'-1}=\frac{1}{(m+1)!}\frac{d^{m+1}\psi}{d\rho^{m+1}}(\rho_h)$$

and these derivatives vanish for $m+1\le m_h-1$ because ρ_h is a root of ψ with multiplicity m_h. □

Corollary 5.1. *When all the eigenvalues ρ_h are simple, the corresponding processes $\boldsymbol{\zeta}_{t,\rho_h}$ in Equation (8) are scalar processes and they are all non-deterministic.*

Proof. It is enough to show that the vector $c=C^{-1}u$ appearing in Equation (3) and at the beginning of Section 3.3 has all its components different from zero. The i-th row of the matrix C^{tr} is $(\rho_i^{r-1},\rho_i^{r-2},\ldots,\rho_i,1)^{\mathrm{tr}}$. Let $a_i(z)=a_{i,1}z^{r-1}+a_{i,2}z^{r-2}+\ldots+a_{i,r}$ denote the polynomial determined by the equalities $a_i(\rho_j)=\mathbf{1}_{i=j}$. Then $C^{\mathrm{tr}}(a_{i,r},a_{i,r-1},\ldots,a_{i,1})^{\mathrm{tr}}=(a_i(\rho_1),a_i(\rho_2),\ldots,a_i(\rho_r))^{\mathrm{tr}}$ has all components equal zero except for the i-th one that equals one. Therefore c satisfies $c_i=c^{\mathrm{tr}}C^{\mathrm{tr}}(a_{i,r},a_{i,r-1},\ldots,a_{i,1})^{\mathrm{tr}}=u^{\mathrm{tr}}(a_{i,r},a_{i,r-1},\ldots,a_{i,1})^{\mathrm{tr}}=a_i(1)\neq 0$, because $a_i(1)=0$ together with $a_i(\rho_j)=0, j\neq i$ would imply that a_i vanishes identically. □

B. Computation of the matrix C^{-1}

Let us introduce the polynomial $\alpha(z) = \sum_{j=1}^{r} \alpha_j z^{r-j}$ and denote $\boldsymbol{\alpha}$ the row vector of its coefficients. The product of $\boldsymbol{\alpha}$ times the j-th column of the h-th block C_h of the matrix C is

$$\boldsymbol{\alpha} \begin{pmatrix} c_{1,j}^{(h)} \\ \vdots \\ c_{r,j}^{(h)} \end{pmatrix} = \sum_{i=1}^{r} \alpha_i c_{i,j}^{(h)} = \sum_{i=1}^{r} \alpha_i \frac{\mathcal{D}^{m_h - j} \rho_h^{r-i}}{(m_h - j)!} = \frac{\mathcal{D}^{m_h - j} \alpha(\rho_h)}{(m_h - j)!}$$

After dividing C^{-1} in adjacent blocks of sizes $m_h \times r$, we find that the elements in the i-th row of the h-th block are the coefficients of the polynomial $\alpha_{h,i}$ of degree $r-1$ with the properties

i. $\mathcal{D}^{m_l - j} l\alpha_{h,i}(\rho_l) = 0$ for $l \neq h$ and $j = 1, 2, \ldots, m_l$,

ii. $\mathcal{D}^{m_h - i} \alpha_{h,i}(\rho_h) = (m_h - i)!$ and

iii. $\mathcal{D}^{m_h - j} \alpha_{h,i}(\rho_h) = 0$ for $1 \leq j \leq m_h, j \neq i$.

From i. it follows that $\alpha_{h,i}(z)$ is divisible by $a_h(z) = \prod_{l \neq h}(z - \rho_l)^{m_l}$. Then $\alpha_{h,i}(z) = a_{h,i}(z) a_h(z)$ with $a_{h,i}$ polynomial of degree at most $m_h - 1$. Now notice that the polynomial $b_{h,i}(z) = (z - \rho_h)^{m_h - i} \frac{a_h(z)}{a_h(\rho_h)}$ satisfies $\mathcal{D}^{m_h - j} b_{h,i}(\rho_h) = 0$ for $j > i$ and $\mathcal{D}^{m_h - i} b_{h,i}(\rho_h) = (m_h - i)!$. This suffices to establish that $a_{h,1} = b_{h,1} = (z - \rho_h)^{m_h - 1} \frac{a_h(z)}{a_h(\rho_h)}$ solves i., ii. and iii. for $i = 1$.

The polynomial $b_{h,2}$ satisfies conditions i., ii. required for $\alpha_{h,2}$ and also iii. for $j > 2$, but $\mathcal{D}^{m_h - 1} b_{h,2}(\rho_h)$ is different from zero, so that iii. is not satisfied for $j = 1$. By subtracting to $b_{h,2}$ a multiple of $\alpha_{h,1}$ the conditions already verified still hold, and the multiple can be chosen so as to have the remaining one verified, namely

$$\alpha_{h,2}(z) = b_{h,2}(z) - \mathcal{D}^{m_h - 1} b_{h,2}(\rho_h) \frac{1}{(m_h - 1)!} \alpha_{h,1}(z)$$

has the required properties.

In general, the recursive formulas

$$\alpha_{h,i}(z) = b_{h,i}(z) - \sum_{j=1}^{i-1} \mathcal{D}^{m_h - j} b_{h,i}(\rho_h) \frac{1}{(m_h - j)!} \alpha_{h,j}(z)$$

permit to obtain the coefficients of the polynomial $\alpha_{h,i}(z)$, which are the elements of the corresponding row of C^{-1}. In the simplest case, viz. when all eigenvalues are simple, then $\alpha_h(z)(= \alpha_{h,1}(z)) = \prod_{j \neq h} \frac{z-\rho_j}{\rho_h - \rho_j}$. In particular, the vector $c = C^{-1}u$ with $u = (1, 0, 0, \ldots, 0)^{\mathrm{tr}}$ is

$$c = \left(\frac{1}{\prod_{j \neq 1}(\rho_1 - \rho_j)}, \frac{1}{\prod_{j \neq 2}(\rho_2 - \rho_j)}, \ldots, \frac{1}{\prod_{j \neq r}(\rho_r - \rho_j)} \right)^{\mathrm{tr}}$$

$$= \left(\frac{1}{\mathcal{D}R(\rho_1)}, \frac{1}{\mathcal{D}R(\rho_2)}, \ldots, \frac{1}{\mathcal{D}R(\rho_r)} \right)^{\mathrm{tr}}$$

with $R(z) = \prod_{j=1}^{r}(z - \rho_r)$.

References

Brockwell, A. E., Brockwell, P. J., 1999. A class of non-embeddable ARMA processes. J. Time Ser. Anal. 20 (5), 483–486.

Brockwell, P. J., 1995. A note on the embedding of discrete–time ARMA processes. J. Time Ser. Anal. 16, 451–460.

Brockwell, P. J., 2004. Representations of continuous time ARMA processes. J. Appl. Probab. 41, 375–382.

Brockwell, P. J., Davis, R., 2002. Introduction to Time Series and Forecasting, 2nd Edition. Springer.

Brockwell, P. J., Hannig, J., 2010. CARMA(p, q) generalized random processes. J. Statistical Planning and Inference 140, 3613–3618.

Chan, K. S., Tong, H., 1987. A note on embedding a discrete parameter ARMA model in a continuous parameter ARMA model. J. Time Ser. Anal. 8, 277–281.

Doob, J. L., 1944. The elementary Gaussian processes. Annals of Mathematical Statistics 15 (3), 229–282.

Gel'fand, I. M., Vilenkin, N. Y., 1964. Generalized Functions. Academic Press.

He, S. W., Wang, J. G., 1989. On embedding a discrete-parameter ARMA model in a continuous-parameter ARMA model. J. Time Ser. Anal. 10, 315–323.

Huzii, M., 2006. Embedding a Gaussian Discrete-time Autoregressive Moving Average process in a Gaussian Continuous time Autoregressive Moving Average process. Journal of Time Series Analysis 28 (4), 498–520.